Assessing Visual Short-Term Memory in 5- to 12-Month-Old Infants Using an Eye-Tracking Change-Localization Task at Set Sizes Three and Four

By

VAN PHAM

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Child Development

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Lisa Oakes, Chair

Katharine Graf Estes

Yuko Munakata

Committee in Charge

2023

Abstract

Previous visual short-term memory research has shown that infants can store visual information in (VSTM) and that the amount of information infants can store changes across development. Recently, there has been a shift toward understanding how infants store information in VSTM. We tested 5- to 12-month-old infants (N = 57, 31 girls) from the Greater Sacramento area of California, USA, in an eye-tracking *change localization task*. Infants saw trials with the following sequence: a 500-ms *sample array* of three or four (set sizes) colored circles, followed by a 300-ms *delay array* with a blank screen, and finally a 2000-ms *test array* in which one circle chosen at random changes color. At both set sizes (3 and 4), infants successfully localized the change and preferred the changed item more than chance. Moreover, we found that when infants fixated the *to-be-changed* item prior to the change onset, they showed a stronger preference for the changed item during the test array compared to when they did not fixate the to-be-changed item. These results add to our growing understanding of the development of VSTM in infancy, and demonstrate the importance of infants' attention on their encoding of information in VSTM and in this task their localization of the change in the test array.

Keywords: Visual short-term memory; Infancy; Eye tracking; Memory; Cognition

Assessing Visual Short-Term Memory in 5- to 12-Month-Old Infants Using an Eye-Tracking Change-Localization Task at Set Sizes Three and Four

1.1 What is visual short-term memory and why is it important?

Visual short-term memory (VSTM) is an important memory system that is used for the temporary storage and maintenance of visual information across interferences that occur during eye movements and blinks (Luck, 2007). Similar to other components of the working memory (WM) system, VSTM is associated with individual differences in cognitive processing (Luck & Vogel, 1997). For example, children with specific language impairment (SLI) appear to have impaired VSTM performance (Leclercq et al., 2021). Likewise, children with Down syndrome, those with attention-deficit/hyperactivity disorder (ADHD), schizophrenia, Alzheimer's Disease, and dementia also have impaired VSTM performance (Chapman & Hesketh, 2001; Coffman et al., 2020, Dovis et al., 2013; Parra et al., 2010).

This memory system appears to emerge in infancy; there is evidence of VSTM in infants as young as 4-months of age. The development of VSTM rapidly increases in the first year of life (Oakes et al., 2006, 2009, 2011; Ross-Sheehy et al., 2003). The emergence of VSTM during infancy plays an important role in learning about object properties and forming categories, searching successfully for objects in the A-not-B task, and discriminating between different displays (Lange-Küttner, 2008; Mix et al., 2002; Oakes et al., 2011). Much of the work on VSTM in infancy has been focused on demonstrating *that* infants can store information in VSTM and how information they store changes across development (Beckner et al., 2020; Cantrell et al., 2019; Eschman & Ross-Sheehy, 2023; Ross-Sheehy & Eschman; 2019). Recently, researchers have turned their attention to understand *how* infants store information in VSTM. Thus, one goal of the present study is to add to our understanding of how infants' VSTM performance is influenced by what information they actually pay attention to.

1.2 How is VSTM studied in infants?

VSTM has been studied during infancy using two different change-detection tasks: *simultaneous stream* (also known as *dual-stream procedure*) and *one-shot* tasks. In the simultaneous stream procedure, infants are shown two stimulus streams, side-by-side, at the same time (Kwon et al., 2014; Oakes et al., 2006, 2011, 2017; Ross-Sheehy et al., 2003). The streams typically include arrays of shapes or objects that briefly appear and disappear repeatedly. In the *changing* stream, one or more items or objects are changed at each reappearance. In the other stream, the *non-changing* stream, none of the items or objects change. Infants' VSTM is inferred from their looking preference to the changing stream compared to the non-changing stream. Although this procedure has been useful for demonstrating VSTM and its development in infancy (Kwon et al., 2014; Oakes et al., 2006, 2011, 2017; Ross-Sheehy et al., 2003), because conclusions are drawn from how long infants look at the arrays as a whole, it is difficult to know from this task whether they detected the specific changes in these arrays.

One study used a variation of this task and suggested that we can sometimes draw conclusions about infants' looking at individual items from their looking at an entire array. Ross-Sheehy et al. (2011) showed infants streams of 3 items that appeared and disappeared and recorded their looking to the streams as a whole, as in the simultaneous streams task. However in this study, infants were presented with only one stream at time. On every trial, one of the three items changed on each reappearance. On half of the trials, the item that was changing *rotated* continuously and on the other trials a non-changing item rotated. 5-month-old infants looked longer at the streams in which the rotating item changed than at the streams in which a non-rotating item changed, suggesting that they primarily attended to the rotating item in these arrays.

More direct evidence that infants notice precisely what changed comes from the use of eye tracking. Eye-tracking procedures allowed the development of a *one-shot change detection task.* In this task, participants are presented with a series of trials that each involve a *sample array* containing a few items (e.g., shapes or objects), followed by a brief blank display (the *delay* period), and lastly a *test array* (Luck & Vogel, 1997). In this test array, one or more of the items may have changed. Although some conclusions have been drawn from infants' looking at changed versus non-changed arrays (Ross-Sheehy et al. 2003), many studies have drawn conclusions about infants' and young children's change detection from their looking to the particular items in the array (Beckner et al., 2020; Cantrell et al., 2019; Eschman & Ross-Sheehy, 2023; Oakes et al., 2017; Ross-Sheehy & Eschman, 2019). That is, what proportion of their looking is directed at the changed item during the test array. Thus, in the *one-shot change detection task*, we can infer VSTM from infants' and young children's localization of the change (i.e., having a preference for the changed item) rather than just overall preference to the whole array.

Studies often vary the number of items in an array to test the limits of VSTM in infants and young children (Buss et al., 2018; Kwon et al., 2014; Eschman & Ross-Sheehy, 2023; Oakes et al., 2011; Perone et al., 2011; Ross-Sheehy & Eschman, 2019; Simmering, 2012). In both procedures (*simultaneous stream procedure* and *one-shot change detection task*), the set size (SS) refers to the number of items or objects shown in the arrays. Participants' VSTM capacity is established by varying the set size and observing whether or not they detect a change. In adults, for example, VSTM capacity is estimated to be four to six items (Luck & Vogel, 1997). It is tempting to conclude that we can establish infants' VSTM capacity by increasing the number of items in the array. However, infants may show preferences for the changing display indicating they have memory for the items shown even though the number of items (set size) shown is above infants' VSTM capacity. This is because infants may be able to detect a change if they have stored a subset of the items (i.e., not all items) shown in their VSTM. Moreover, using one-shot change detection tasks, Buss et al. (2018) and Simmering (2012) found that toddlers' VSTM appears to be more limited compared to previous reports of infants' VSTM performance.

1.3 Previous VSTM findings

Estimating set size in infancy is complicated. Increasing set sizes and finding that infants showed a preference for the changing display in the simultaneous stream procedure does not necessarily mean that we can conclude what infants' VSTM capacity is. As previously discussed, infants may show a preference for a changing display even when they have a smaller VSTM capacity than the number of items in the arrays. That is, they may only store some items in the array and detect a change if one of those stored items changes. In addition, infants may chunk information (i.e., recognizing temporal regularities in the task) (Kibbe & Feigenson, 2016; Moher et al., 2012), and as a result detect changes in arrays that are larger than their VSTM capacity.

In general, studies using simultaneous stream procedures have shown that infants 6 months and younger do not detect changes in arrays of 2 or more items (Oakes et al., 2006; 2009; Ross-Sheehy et al., 2003). This has been taken as evidence that younger infants have difficulty with multi-item arrays, perhaps because they are unable to bind object identity to location (Oakes et al., 2006). However, research using the one-shot task has yielded more inconsistent results. Oakes et al. (2013) used the one-shot task with 6- and 8-month-old infants. In this study, 8-

month-old infants detected changes at set size 2, but 6-month-old infants did not. Ross-Sheehy and Eschman (2019), in contrast, found in a one-shot change detection task that 5- and 8-month-old infants detected a change at set size 2, and 11-month-old infants detected a change at set sizes 2 and 3. This suggests clear evidence of VSTM performance improvement from 5- to 11-months of age, although the results varied somewhat.

This research has raised questions about *how* infants encode information in VSTM. One possibility is that infants can encode in VSTM items from multiple item arrays if they actually look at and attend to those items. Recall that Ross-Sheehy et al. (2011) showed infants one stream at a time instead of simultaneous streams (i.e., known as dual streams; one stream on each side of the screen). They found that 5-month-old infants looked longer at changing streams with 3-item arrays in which the changing item continuously rotated than at changing streams in which a non-changing item continuously rotated. This finding suggested that young infants can encode salient items into their memory.

In support of this conclusion, several studies using the one-shot task have shown that young infants seem to store in VSTM items they fixate during the sample array. Cantrell et al. (2019) found that 6-month-old infants showed evidence of VSTM in 2-item arrays if they fixated on the ‘to be changed’ item during the sample array (i.e., the item that is later changed in the test array). Beckner et al. (2020) induced infants to look at an item in a 2-item array during the sample period by rotating it while keeping the other item stationary. During the test array, the rotated item in the sample array stayed stationary, but it either changed or did not change in color. They found that although 4-month-old infants looked equally at the rotated item when it changed and did not change color during the test array, suggesting that they did not encode the fixated item (i.e., rotated item) into VSTM, 8.5-month-old infants did look longer at the rotated item during the test array when it changed color versus when it did not change color.

Ross-Sheehy and Eschman (2019) examined whether across age, infants would detect changes in increasingly larger set sizes in the one-shot task, and how their fixation of the to-be-changed items contributed to that change detection. In this study, Ross-Sheehy and Eschman (2019) tested 5-, 8-, 11-month-old infants, and adults in a one-shot change detection task at set sizes 2, 3, and 4. Infants at all ages showed a preference for the changed item at set size 2, but only 11-month-old infants showed a change preference at set size 3. Adults showed a change preference for all three set sizes (2-4). Moreover, when collapsing across set sizes, Ross-Sheehy and Eschman (2019) found that at all ages (including adults), participants only showed a change preference when they had fixated the to-be-changed item during the sample array. However, because they did not analyze this separately for each set size, their data do not allow conclusions about how the fixations during the sample array are related to change detection at the different set sizes. Examining differences between set sizes is one goal of the current study.

1.4 The current study

The current study seeks to bridge the gap in understanding VSTM capacity in infancy by using an eye-tracking change localization task to test a wide range of ages in infancy with two set sizes. In this study, we tested a group of 5- to 12-month-old infants in a change localization task at set sizes 3 and 4. We are interested in whether 5- to 12-month-old infants show VSTM performance at each of the set sizes and whether or not there are age differences. The timing of the phases in our change localization task (i.e., 500-ms sample array, 300-ms delay/retention phase, and 2000-ms test array) are similar to previous studies (Beckner et al., 2020; Cantrell et al., 2019; Oakes et al., 2013), using larger set sizes (3 and 4) and arrays that are more similar to the studies by Eschman and Ross-Sheehy (2023; Ross-Sheehy & Eschman, 2019). Like Ross-Sheehy and Eschman (2019), we tested all our infants at both set sizes 3 and 4, allowing us to make within-subject comparisons of the different set sizes. Our primary goal was to understand *how* infants store information in VSTM. In previous work (Beckner et al., 2020; Cantrell et al., 2019; Ross-Sheehy & Eschman, 2019), researchers asked how fixations during the sample array relates to change detection. Thus, another goal of this study is to examine how infants' VSTM performance is influenced by what they looked at during the sample and delay period. Based on previous findings, we are interested in specifically how infants look at the to-be-changed item during the sample array and delay period and how that is related to change preference/localization during the test array.

Method

2.1 Participants

The data for this study were previously collected as part of a larger study. The final sample for the present analyses included 57 5- to 12-month-old infants *(M* = 265.0 days, *SD* = 67.6 days); all of the infants will be included in most of the analyses, but only 52 infants will be included in our overall change localization score analyses (see Data Processing section for more details). Thirty-one of the 57 infants were girls and 26 were boys. Infants in our study came from different racial backgrounds, including White (n = 41), multi-racial (n = 5), and Asian or Asian American (n = 5); the racial background of 6 infants was not reported. Our sample also included diverse ethnic backgrounds; 14 of our infants were identified as Hispanic. Most of the mothers in our sample had earned a high school diploma (98%; n = 56) and 74% (n = 42) of the mothers earned at least a 4-year college degree (i.e., Bachelor's degree or higher).

Infants were recruited from a pool of potential research participants at the Center for Mind and Brain (CMB). The names of infants were obtained from the State of California Office of Vital Records. Families who had a home address within 30 miles of the CMB located in Davis, California were sent flyers with information about how to participate in studies at the CMB . When children approached the age we were testing in the study, they were contacted via email or phone and were provided information regarding the study. If they were interested, a screening procedure was conducted to determine whether or not infants met the criterion for our study (e.g., no vision or hearing problems, medical or neurological problems, or not born prematurely). All eligible infants were scheduled for an appointment and those who participated received a small t-shirt or a toy valuing no more than $10.

An additional 17 infants were tested but excluded from our final sample due to being born prematurely (n = 2), having family history of colorblindness that put them at higher risk for colorblindness (e.g., boys who had maternal uncles who were colorblind, n = 3), an inability to calibrate their eyes to the eye tracker (n = 6), experimenter error (n = 1), or being too fussy to contribute any data to the study (n = 5).

2.2 Apparatus

The data for this study was collected using a Sensomotoric Instrument (SMI)-Red eye tracker and a Dell laptop provided by SMI. The eye tracker recorded eye movement at a rate of 120 Hz. The eye tracker used an infrared light source to determine point-of-gaze (POG) using pupil and corneal reflection. The eye tracker was attached to a 22 inch (1680 by 1050 pixels) LCD monitor, which was attached to an ergo arm that allowed the monitor and eye tracker to be moved and positioned for the optimal detection of each infant's eyes. In addition, a Logitech Carl Zeiss Tessar 2.0/3.7 2MP Autofocus webcam attached to the top of the monitor to record and display the infant's behavior to the experimenter. To minimize distractions to the infant, a large white cloth screen was placed behind the monitor and eye tracker to block the infant's view of the additional equipment (e.g., Dell laptop, speaker system).

2.3 Stimuli

The experimental stimuli were arrays of 3 or 4 differently colored circles. The circles were one of 8 highly discriminable colors (e.g., green, orange, purple, red, cyan, yellow, black, brown), and were approximately 4.13 cm in diameter (3.94° at a viewing distance of 60 cm). The circles were arranged into one of two different configurations (i.e., orientations) for each of the set sizes. This helped to eliminate any biases to any particular location on the screen. The configurations involved the circles being evenly distributed around the center, so a triangle at set size three and a square at set size 4. The difference in the orientations can be seen in *Figure A*. The center-to-center distances of the circles was 14 cm for set size 3 (13.31°) and 11.5 cm for set size 4 (10.95°). The eccentricity (i.e., edge to edge distance) of the circles was 17.5 cm for set size 3 (16.6°) and 15 cm for set size 4 (14.25°). In addition, we presented brief cartoon clips (e.g., Sesame Street, Blue's Clues, Alice in Wonderland) at the beginning of each VSTM block to keep infants engaged in the task.

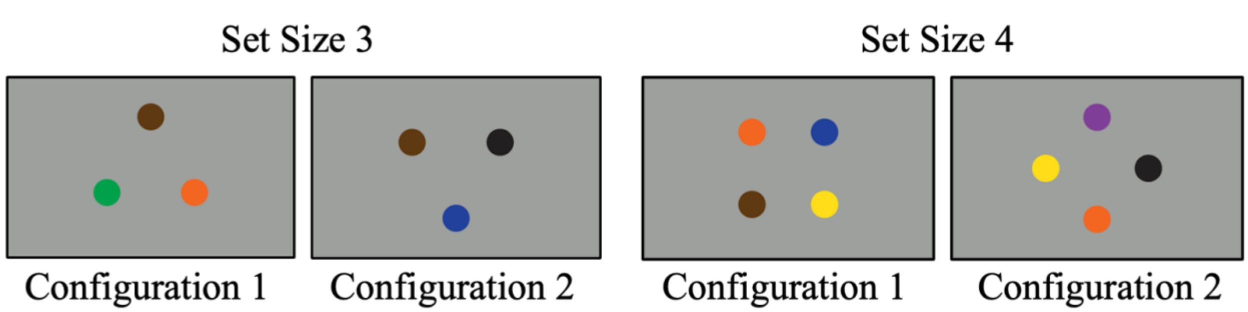


Figure A. Set Sizes Three and Four Configurations

2.4 Procedure

The procedures of this study were approved by the Institutional Review Boards (IRB) of the State of California and the University of California, Davis (UC Davis). Written informed consent (i.e., signed consent forms approved by the State of California and UC Davis IRB) were obtained from parents prior to the start of the study after experimenters read and explained the procedures.

The data analyzed in this study were part of a larger study involving several visual cognitive tasks, including the VSTM task reported here. This task occurred toward the end of a session that lasted an hour and a half to two hours, and after multiple eye-tracking studies and a motor assessment. Only data from the VSTM task will be analyzed in this study.

Infants were seated in a highchair or parent’s lap in a sound attenuated testing room and were seated approximately 60 cm from the monitor and eye tracker. During the eye-tracking study, parents were asked to wear felt-covered sunglasses to prevent them from looking at the screen or biasing the infant's looking behavior.

Before the start of the study, the experimenter used SMI iView software to detect the infant’s eyes and move the ergo arm of the monitor when necessary. A child-appropriate video was shown during this process to keep infants entertained. After the eyes were detected, the experimenter started the SMI automatic calibration process, which involves 5 looming circle stimuli (5-point calibration) that were visually entertaining to infants. After the calibration process, a validation process occurred and images of a yellow rubber ducky and a chirping sound were presented in each of the four corners of the screen. The SMI software gave the experimenter feedback about the accuracy of the point of gaze detection during this validation process. At that point the experimenter chose to either redo the calibration (if the infant’s gaze was greater than 2° degree of the validation locations) or continue with the experiment (if the infant's gaze was within 2° degree).

Before the start of each VSTM trial, a blinking fixation cross (3.34˚ h × 2.86˚ w) accompanied by a sound (i.e., ringing sound) was presented until infants looked at the center of the screen. Once the SMI detected that the infants fixated the cross for 200 ms, the trial began. Infants were shown trials with the following sequence: a 500-ms *sample array* of three or four (set sizes) colored circles, followed by a 300-ms *delay array* with a blank screen, and finally a 2000-ms *test array* in which one randomly selected circle changed color from the sample array. This process repeated for a total of 21 trials at set size 3 and 20 trials at set size 4; see *Figure B* for a schematic of the trial sequence. There were a total of 7 VSTM blocks: six blocks of 6-trials, and one block of 5-trials.

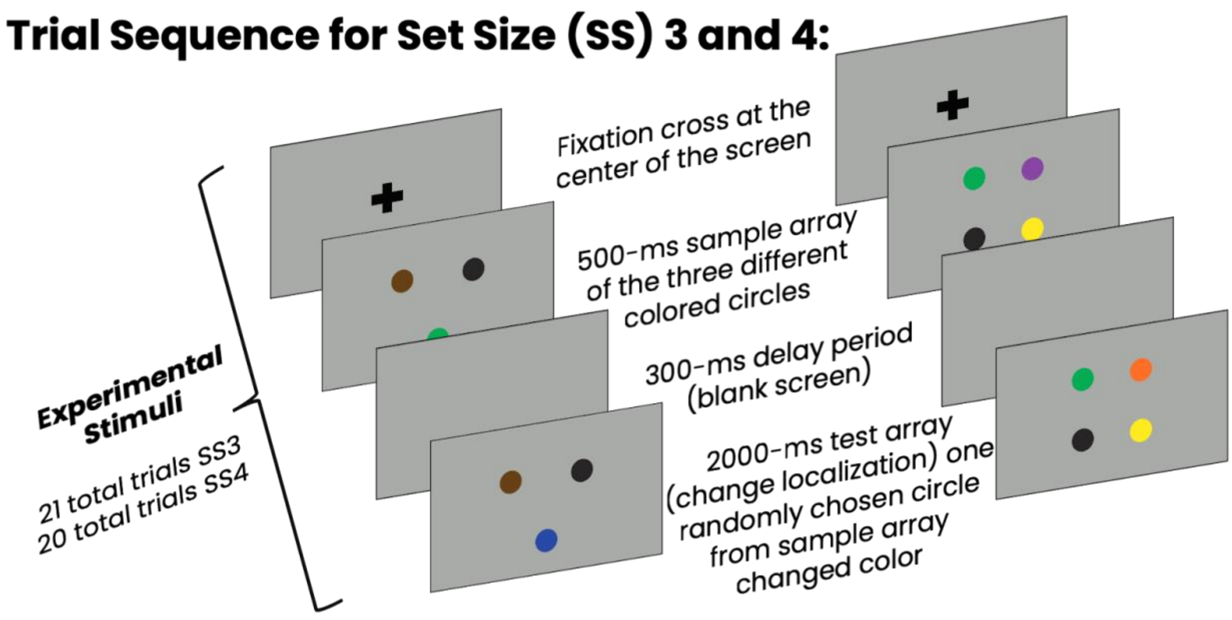


Figure B. Trial Sequence for Set Sizes Three and Four

2.5 Data Processing

We generated several measures from each trial. During the sample and delay period (i.e., the *prechange period*), we calculated (1) total looking to the entire array, (2) the amount of looking to the to-be-changed item or item location, and (3) the amount of looking to the locations of the items that remain unchanged after test. During the test array (i.e., the *postchange period*), we calculated (1) the total amount of looking to all of the three or four items in the array, and (2) *change localization score*, which is the proportion of time infants spent looking at the changed item/total looking at all three or four items (depending on which set size we are calculating the change localization score for). The analysis window for the test array included only fixations that happened 100 ms after the onset of the test array to avoid including any fixations that happened before the onset of the test (change) array since it takes 100-200 ms to plan a saccade (Hyun et al., 2009; Oakes et al., 2017).

We required that infants look at the *prechange period* for at least 100 ms and at the *postchange period* for at least 200 ms for any trial to be included in the analyses. These inclusion criteria were based on previous work by Oakes et al. (2013, 2017) and Beckner et al. (2020). There were 96 trials (out of a total of 1085 trials) that were excluded using these criteria.

Infants contributed an average of 8.8 trials (SD = 5.6) for set size 3 and an average of 9.0 trials (SD = 5.5) for set size 4. Infants had to have usable data on a minimum of two trials at each set size to be included in overall change localization score analyses for each set size (see section *2.6 Analysis Plan* for more information), but we included all infants who had data on at least 1 trial for all the other analyses. There were 5 infants (out of 57 infants) for each set size that were excluded from not providing enough usable data for overall change localization score analyses.

2.6 Analysis Plan

The data were analyzed in several steps to address several different research questions.

1. *Are there any differences in infants' looking behavior for the two set sizes, three and four, during the prechange and postchange periods?*

Our first analyses simply examined any differences in infants' looking behavior between the two set sizes. All of the descriptives were conducted for both the *prechange* and *postchange* periods and for both set sizes. First, we used BeGaze, SMI's data analysis software, to create areas of interest (AOIs) for each item in our arrays. Our AOIs were approximately 6.8 cm tall by 6.8 cm wide (approximately 6.5˚ h x 6.5˚ w at a viewing distance of 60 cm). These AOIs were bigger than our experimental stimuli to account for any calibration inaccuracy. Next, we extracted two measures from BeGaze: (1) *standard fixation parameters* and (2) *the XY coordinates of the eye gaze*. The first measure, *standard fixation parameters*, defines fixations as any period of eye gaze that had at least 80 ms in fixation duration within a dispersion region of 100 pixels. And the second measure, *the XY coordinates*, provided us information on whether the fixations were within one of our AOIs or outside of our AOIs. These two measures provided a data file that included information regarding when each fixation started and ended, the duration of the fixation, the XY coordinates of the fixation, and if the fixation fell within one of our AOIs. From this export, we calculated the total looking during the *prechange* and the *postchange periods* by summing any fixation durations made to any of the areas of interest (i.e., looking at all three or four items depending on set size). We also calculated the total looking time and number of fixations on each trial. Finally, we calculated the number of unique items sampled (i.e., multiple fixations made to the same item would count as one unique item sampled) during the *prechange* and *postchange periods*. We used t-tests and Pearson's correlations to examine any differences in set size or across age for these variables.

2. *Will infants show VSTM for set sizes three and four?*

Our second question was whether infants' change detection differed from chance. This allows us to determine whether infants, overall, stored the items in the arrays into VSTM and detected a change when it occurred. On each trial, we calculated a *change localization* score by dividing the duration of looking at the changed item by the total of looking at all the items in the test array. We then created a single score for each infant at each set size by averaging their change localization scores across trials. For this set of analyses, it was important that a minimum trial criterion was applied because we are using average change localization scores. As a result, we imposed the inclusion criterion described in *Section 2.5 Data Processing*.

Formula for Change Localization Score:

$$Change\ Localization\ Score\ = \frac{Looking\ Time\ to\ the\ Changed\ Item}{Total\ Looking\ Time\ to\ All\ Three\ or\ Four\ Items\ (depends\ on\ set\ size)}$$

We conducted two two-tailed one-sample *t-tests* (one for each set size) to test if infants' average change localization scores (how long infants looked at the change item) across all completed trials are significantly different from chance. Chance is determined by the number of items shown (i.e., set sizes); chance is .33 for set size 3 and .25 for set size 4 because if children did not prefer any particular item, we would expect equal looking at all items shown. Correlations between infant's age and their change localization scores for each set size were also conducted.

3. *Is infants' VSTM performance influenced by what they looked at during the prechange period? That is, are there differences in VSTM performance when infants look at the "to-be-changed" item during the prechange period versus when they did not?*

Next, we examined if infants in our study showed a difference in VSTM performance when they looked at the to-be-changed item during the *prechange period* versus when they did not. We first classified trials in which infants had at least one fixation to the AOI of the to-be-changed item during the *prechange period* as "fixated" and trials in which infants did not make any fixations to the AOI of the to-be-changed item as "not fixated". Then, we calculated each participant's average change localization scores for each trial classification (trials in which infants did look or did not look at the to-be-changed item) for each set size, and compared those trials to chance using two-tailed one-sample *t*-tests.

4. *The use of Trial-Level Linear Mixed-Effects (LME) Models to determine what variables predict infants' change localization score on a trial-level?*

We fit separate LME models for each set size on infants' trial-level change localization scores to ask the question of how infants' change detection varied as a function of age, whether or not they fixated the to-be-changed item during the *prechange period* (trial classification), and trial index. Our trial classification variable was logistic thus *Yes* denotes that infants did fixate the to-be-changed item for that trial and *No* denotes that infants did not fixate the to-be-changed item. In the models, we also included a random intercept of participant ID that denotes the unique ID for each participant. A separate LME model was fitted for each set size because of the differences in the chance level, .33 for set size 3 and .25 for set size 4.

Model Fitted for Both Set Sizes:

lmer(Change Localization Score ~ scale(Age in Days)*Trial Classification (Yes vs. No)*Trial Index + (1|Participant))

Results

We analyzed the data in several stages. First, we provide descriptive statistics and comparisons of infants' general looking behavior as a function of set size for the two periods: prechange and postchange period. These analyses provide insight into whether infants were equally attentive during the trials of both set sizes and help to rule out the possibility that any differences observed are due to differences in attention. For these analyses, we included any trial in which infants had at least 100 ms of looking during the prechange period and at least 200 ms of total looking during the postchange period (other inclusion criteria were used for other analyses).

Next, we analyzed infants' overall change localization compared to chance. We then examined infants' change localization as a function of what infants fixated during the prechange period. That is, these analyses asked whether or not trials in which infants fixated the to-be-changed item during the prechange period resulted in a higher change localization than on trials in which infants did not fixate the to-be-changed item during the prechange period. To be included in these analyses, all trial-level looking behavior during the prechange (at least 100 ms of looking) and postchange (at least 200 ms of total looking) periods are required and infants must have contributed at least 2 trials for each set size.

Lastly, we fit infants' change localization scores to LMEs to examine the combined effects of age, whether or not infants fixated the to-be-changed item during the prechange period, and trial index on infants' responses to the change. We included in these analyses all trials that met our trial-level criteria (at least 100 ms of looking during prechange and at least 200 ms of total looking during post-change).

3.1.1 General looking behavior during the prechange period

Our first set of analyses provides an overview of infants' attention and gaze behavior during the prechange period, or the 800-ms sample-plus-delay. During this period, infants spent an average of 314.7 ms (*SD* = 66.0 ms) looking at all the AOIs combined for the items in the array for set size 3 and an average of 319.2 ms (*SD* = 66.0 ms) looking at all the AOIs combined for the items in the array for set size 4. Comparison of the looking duration for the two set sizes using paired two-tailed *t*-test revealed that these looking durations did not differ statistically, $t(51) = -0.67$, $p = .51$, $d = 0.09$. Infants made an average of 2.6 fixations (*SD* = 0.3 fixation) per trial for set size 3 and an average of 2.6 fixations (*SD* = 0.4 fixation) per trial for set size 4. The difference in the number of fixations made in the two set sizes did not differ statistically, $t(53) = -0.60$, $p = .55$, $d = 0.08$. The average duration for individual fixations was 250.6 ms (*SD* = 52.0 ms) for set size 3 and 259.1 ms (*SD* = 52.0 ms) for set size 4. Paired two-tailed *t*-test revealed that the average duration for individual fixations for the two set sizes did not differ significantly, $t(53) = -1.84$, $p = .07$, $d = .25$. Lastly, infants only sampled an average of 1 unique item (i.e., directed one or more fixations to an individual circle in the array) per trial (*SD* = 0.09 item) for set size 3 and an average of 1.1 item per trial (SD = 0.2 item) for set size 4. Paired two-tailed *t*-test revealed that the number of items sampled during the prechange period for the two set sizes did not differ significantly, $t(51) = -0.73$, $p = .47$, $d = .10$.

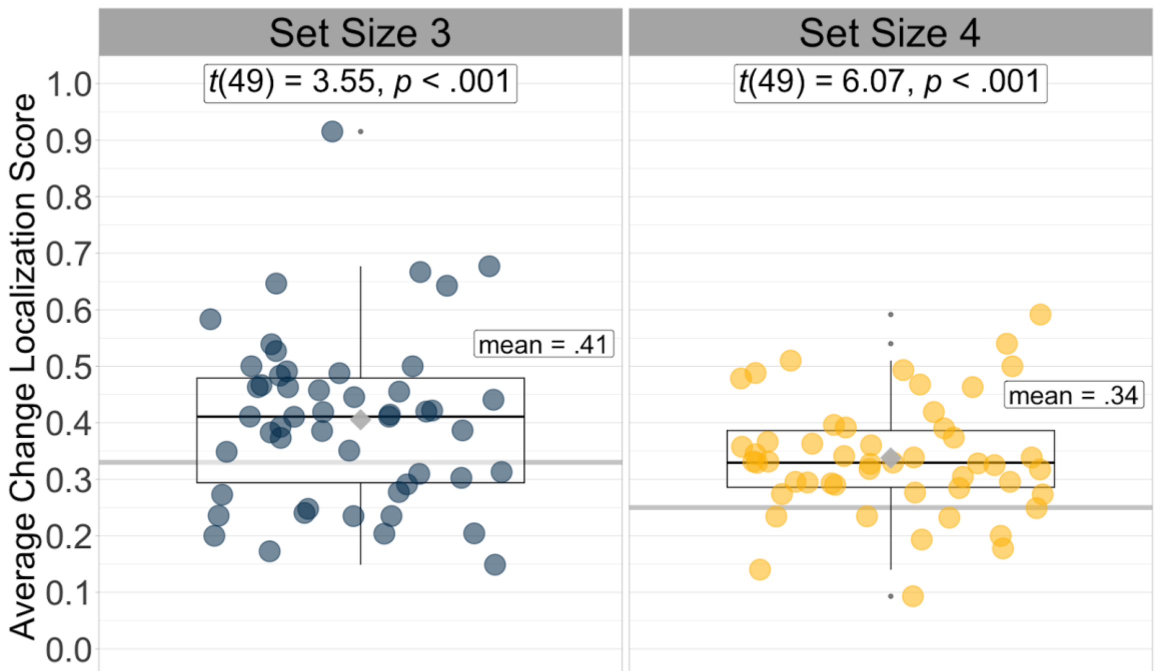
3.1.2 General looking behavior during the postchange period

Our next set of analyses provide an overview of infants' attention and gaze behavior during the postchange period, or the 2000-ms test array (array in which one of the circles changed color). During the postchange period, infants spent an average of 1096.5 ms (*SD* = 293.5 ms) looking towards all the AOIs combined for set size 3 and an average of 1135.8 ms (*SD* = 314.2 ms) looking towards all the AOIs combined for set size 4. Comparison of the looking duration for the two set sizes using paired two-tailed *t*-tests revealed that these looking durations did not differ statistically, $t(53) = -1.82$, $p = .08$, $d = .25$. Infants made an average of 3.4 fixations (*SD* = 1.0 fixations) for set size 3 and 3.5 fixations (*SD* = 1.0) for set size 4. Paired two-tailed *t*-test revealed that the average number of fixations made for the two set sizes did not differ significantly, $t(53) = -0.91$, $p = .37$, $d = .12$. The average fixation duration was 377.1 ms (*SD* = 160.1 ms) for set size 3 and 380.0 ms (*SD* = 130.0 ms) for set size 4. Paired two-tailed *t*-test revealed that the average fixation duration for the two set sizes did not differ significantly, $t(53) = -0.46$, $p = .65$, $d = .06$. Lastly, infants only sampled an average of 1.9 items (*SD* = 0.49 item) per trial for set size 3 and an average of 1.9 items (*SD* = 0.41 item) per trial for set size 4. Paired two-tailed *t*-test revealed that the number of items sampled for the two set sizes did not differ significantly, $t(53) = 0.49$, $p = .62$, $d = .07$.

3.2 Infants' overall change localization compared to chance

Our next analysis was running a two-tailed one-sample *t*-test for each set size compared to chance (.33 for set size 3 and .25 for set size 4; see *Figure C* gray horizontal lines). Before running the *t*-tests, we created an average change localization score at each set size for each infant by averaging the change localization score across all trials they saw at each set size (each single circle on *Figure C* represents one participants' average change localization score). The mean average change localization score for set size 3 was 0.41 (*SD* = 0.15) and for set size 4 was 0.34 (*SD* = 0.10). Overall, infants preferred to look at the changed item more than chance (.33 for set size 3 and .25 for set size 4) in the test array for both set sizes, $t(49) = 3.55$, $p < .001$, $d = .50$

(set size 3) and *t*(49) = 6.07, *p* < .001, *d* = .86 (set size 4). We correlated infants' overall change localization with infant age in days and found no significant correlation between age and change localization for both set sizes, *r*(48) = .17, p = .24 (set size 3) and *r*(48) = .13, *p* = .37 (set size 4).



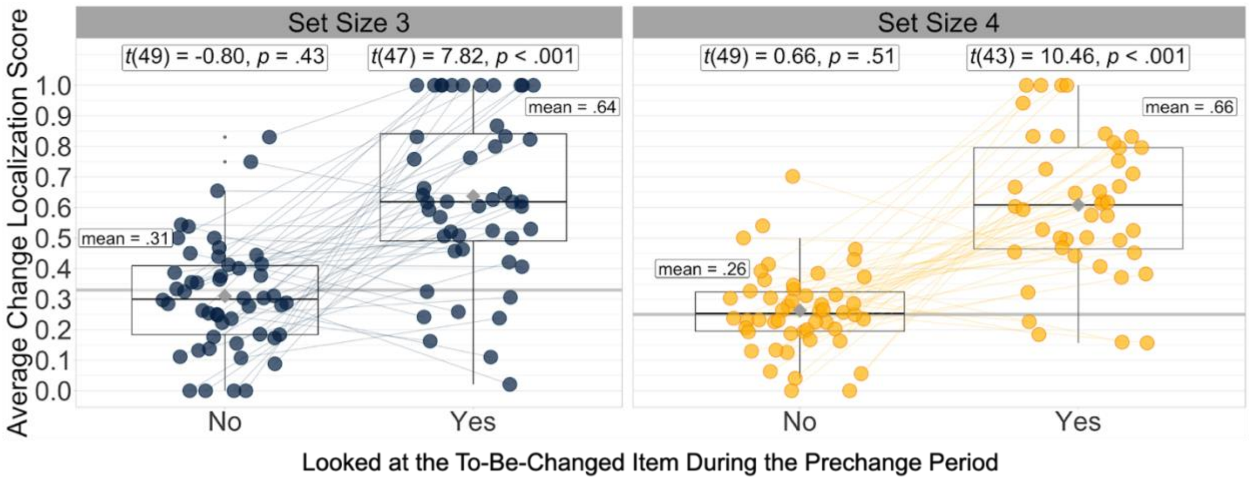
Note. Each circle represents a single participant's average change localization score across all the trials they completed for set sizes (SS) 3 (left, blue dots) and 4 (right, yellow dots). Light gray horizontal lines represent the chance line for each set (.33 for SS3 and .25 for SS4). Light gray diamond shapes denote the mean change localization score for the sample as a whole for each set size. Bold black line within the boxplots indicates the median change localization score. Lines extending from the bottom of boxplots are the lower quartile (Q1) and lines extending from the top of the boxplots are the upper quartile (Q3); calculated using the formula 1.5*IQR. Smaller gray circles are the extreme data points from the *boxplot* function in *R* (R Core Team, 2019).

Figure C. Overall Change Localization Scores for Set Size Three and Four

3.3.2 Infants' overall change localization for trials where they did and did not look at the to-be-changed item

Previous studies have shown that infants only prefer a change when they fixate the to-be-changed-item during the pre-change period (Beckner et al., 2020; Cantrell et al., 2019; Eschman & Ross-Sheehy, 2023; Ross-Sheehy & Eschman, 2019). In our sample, infants fixated the to-be-changed item during the prechange period on about 28% of the trials (135 out of 478 trials) at set size 3 and about 20% of the trials (97 out of 495 trials) at set size 4. On average, we would expect that infants would fixate the to-be-changed item during the prechange period on about 33% of trials for set size 3 and about 25% of trials for set size 4. This is because each item has a 33% chance for set size 3 or 25% chance for set size 4 of changing at random.

To ask whether infants showed a stronger change localization on trials in which infants fixated the to-be-changed item during the prechange period, we calculated the average change localization scores for those trials in which infants did fixate the to-be-changed item during the prechange period (the *Yes* group in *Figure D*) and those trials in which infants did not fixate the to-be-changed item during the prechange period (the *No* group in *Figure D*). Comparison of these scores to chance revealed that infants preferred to look at the changed item more than expected by chance for both set sizes only when they fixated the to-be-changed item during the prechange period, set size 3, $t(47) = 7.82$, $p < .001$, $d = 1.13$, and set size 4, $t(43) = 10.46$, $p < .001$, $d = 1.58$. When infants did not fixate the to-be-changed item during the prechange period, they did not prefer the changed item more than expected by chance, $t(49) = -0.80$, $p = .43$, $d = 0.11$ for set size 3 and $t(49) = 0.66$, $p = .51$, $d = 0.09$ for set size 4.



Note. Each circle represents a single participant's average change localization score across all trials for each trial type: Yes (did look at the to-be-changed item during the prechange period) and No (did not look at the to-be-changed item during the prechange period). Connecting thinner lines between each circle represents data from the same infant (some infants do not have both trial types). Light gray horizontal lines represent the chance line for each set (.33 for SS3 and .25 for SS4). Light gray diamond shapes denote the mean change localization score for the respective set sizes and trial types. Bold black line within the boxplots indicates the median change localization score. Lines extending from the bottom of boxplots are the lower quartile (Q1) and lines extending from the top of the boxplots are the upper quartile (Q3); calculated using the formula 1.5*IQR. Smaller gray circles are the extreme data points from the *boxplot* function in *R* (R Core Team, 2019).

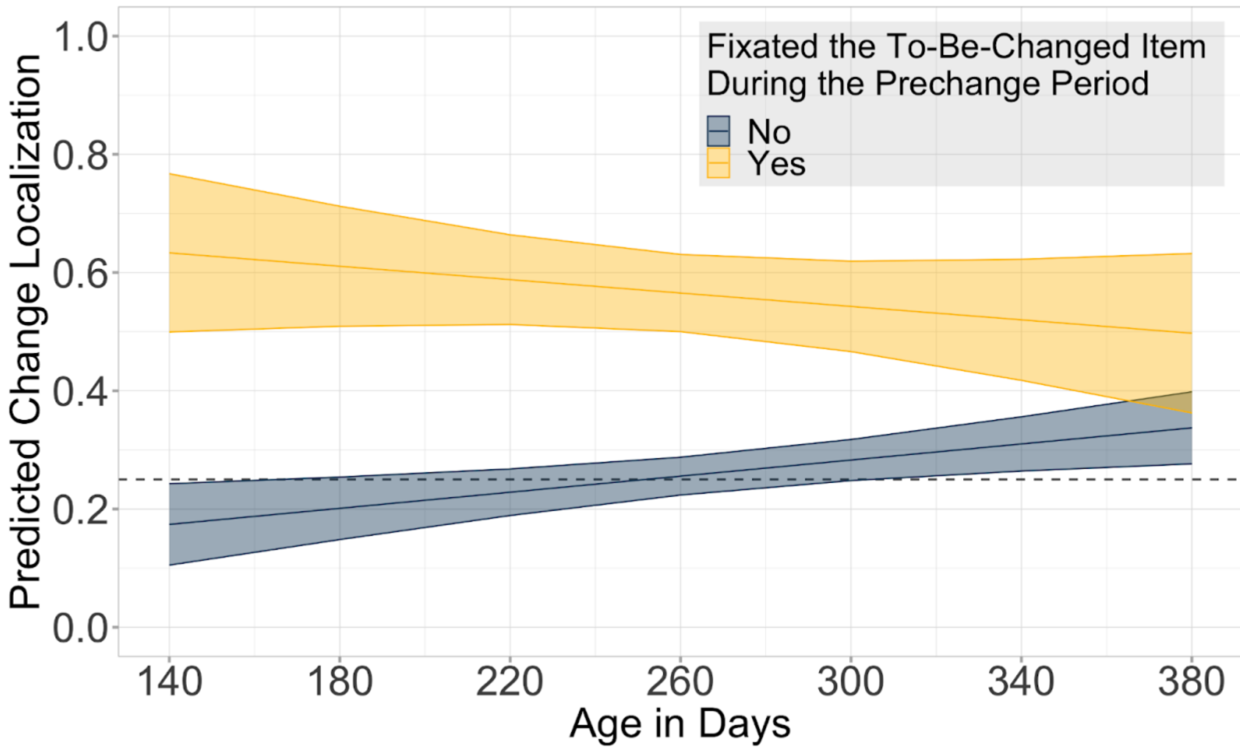
Figure D. Overall Change Localization Scores for Set Size Three and Four

3.4 Trial Level Linear Mixed-Effects (LME) Models for Infants' Change Localization at Set Sizes 3 and 4

Finally, we fit LME models separately for each set size (described in *Section 2.6 Data Analysis*) to ask how infants' change localization changed as a function of age, whether or not they fixated the to-be-changed item during the *prechange period*, and trial index. We ran a separate LME model for each set size (3 and 4). We fit a separate model for the change localization score for each size because of the difference in chance level at the two set sizes (.33 for set size 3 and .25 for set size 4).

The results of the LMEs showed that infants' change localization at both set sizes significantly differed as a function of whether or not they fixated the to-be-changed item during the prechange period, $t(478) = 5.09$, $p < .001$ (set size 3) and and $t(495) = 4.32$, $p < .001$ (set size 4). Inspection of *Figure C* shows that infants' change localization was significantly greater when they fixated the to-be-changed item during the prechange period compared to when they did not fixate the to-be-changed item. The models also revealed that infants' change localization did not significantly change as a function of age for both set sizes, $t(478) = 0.02$, $p = .98$ (set size 3) and $t(495) = 0.40$, $p = .69$ (set size 4). Although there was no main effect of trial index for set size 3, $t(478) = 0.13$, $p = .90$, trial index was significant for set size 4, $t(495) = -3.73$, $p < .001$. This unpredicted effect was relatively small and showed that infants' decreased their preference for the change item somewhat as the session progressed. However, the change localization remained above chance across the session, so this effect may be due to noise, such as infants becoming less interested in the task as trial progresses.

Additionally, infants' change localization did not significantly change as a function of the interaction between age and whether or not infants fixated the to-be-changed item during the prechange period for set size 3, $t(478) = -1.0$, $p = .32$. Interestingly, infants' change localization did significantly change as a function of the interaction between age and whether or not infants fixated the to-be-changed item during the prechange period for set size 4, $t(495) = -3.73$, $p = .02$. On the set size 4 trials, older infants showed a change localization for the changed item even when they did not fixate the to-be-changed item during the prechange period, whereas younger infants only showed a change localization when they fixated the to-be-changed item (see *Figure E*). Lastly, infant's change localization did not significantly change as a function of the three-way interaction (between infants' age, whether or not infants fixated the to-be-changed item, and trial index) for both set sizes, $t(478) = -1.2$, $p = .24$ (set size 3) and $t(495) = 0.98$, $p = .33$ (set size 4).



Note. Confidence intervals (shaded areas) were calculated using the *emmeans* (estimated marginal means) package in *R* from the LME model listed above for set size 4. Horizontal dashed line represents the chance (0.25) line for set size 4. Legend colors represent whether or not infants fixated the to-be-changed item during the prechange period.

Figure E. Set Size 4 Predicted Change Localization as a Function of Age and if Infants Fixated the To-Be-Changed Item During the Prechange Period

Discussion

In this study, we found that 5- to 12-month-old infants successfully localized the change at both set sizes three and four in a VSTM *one-shot* change localization task. Consistent with previous studies (Beckner et al., 2020; Cantrell et al., 2019; Eschman & Ross-Sheehy, 2023; Ross-Sheehy & Eschman, 2019), we found that infants' memory for the changed item during the postchange period heavily depended on whether or not infants fixated the to-be-changed item during the prechange period. On trials in which infants fixated the to-be-changed item during the prechange period, infants showed a stronger preference for the changed item compared to trials in which they did not fixate the to-be-changed item. Thus, in general, our results confirmed the findings from previous literature in this area.

4.1 Demonstrating that infants' can store information in VSTM

In recent years, VSTM research in infancy has shifted from demonstrating *that* infants have VSTM to understanding *how* infants store information in VSTM. Previously, the focus was on establishing at what set sizes infants of different ages detected a change in VSTM tasks. Work using the simultaneous streams procedure, for example, showed that young infants (4- and 6.5-month-old infants) preferred the changing stream only at set size 1, and older infants (10- and 13-month-old infants) preferred the changing stream at set sizes 1, 2, and 3 (Ross-Sheehy et al., 2003). Other research also suggested young infants' VSTM was limited. Using a simultaneous stream task, Kwon et al. (2014) found that 6-month-old infants only preferred the changing stream of arrays of complex objects at set size 1. Even when exposure time to each array was increased. Oakes et al. (2011) showed in the simultaneous stream task that 6-month-old infants had memory for a single location, and only when the location was easily identified by a salient landmark (i.e., reference frame). Similarly, Káldy and Leslie (2005) similarly suggested that 6.5-month-old infants can store a single object representation in a paradigm in which two objects were hidden behind occluders. These studies have confirmed that this is a *memory* problem; young infants tested with multiple-item arrays in streams that did not involve a delay (so one item spontaneously changed to another item) preferred these changing streams (Ross-Sheehy et al.., 2003).

After 6 months, however, infants do detect changes in arrays with multiple objects. Kwon et al. (2014) found that 8-month-old infants preferred the changing stream for both set sizes 1 and 2 even when the arrays involved complex objects, and Oakes et al. (2011) found that 8- and 12.5-month-old infants showed memory for location at set sizes 1, 2, and 3. Several studies using a one-shot task have shown that infants older than 6 months detect changes in arrays of 2 or 3 items (Oakes et al., 2013; Ross-Sheehy et al. citations).

4.2 What is VSTM in infancy?

The shift to understanding how infants store information in their VSTM starts with understanding how VSTM works in infancy, and how it is related to visual attention processes. Mitsven et al. (2018) asked how visual attention is guided by infants' VSTM by using a *cued visual search task*, which is a modified one-shot detection task. In this modification, infants see a brief sample array of a single item. After a brief delay, a test array of two items is presented; one item in the test array matched the item in the sample array (i.e., the *matching item*), whereas the other item in the test array did not match the item in the sample array (i.e., the *non-matching item*. Mitsven et al. (2018) found that 10-month-old infants looked more at the non-matching item than the matching item in the test array. This demonstrated that infants used what was previously stored in their VSTM to guide their subsequent looking behavior (i.e., looking longer at the novel item in the test array because the matching item was previously seen in the sample array).

The Mitsven et al. (2018) study shows how VSTM influences infants' attention. The next question is how does the information infants pay attention to influence their VSTM? Ross-Sheehy et al. (2011) used relative motion cues (i.e., rotation) to induce infants to look at a cued item. As described earlier, Ross-Sheehy et al. presented 5- and 10-month-old infants with streams of three items that repeatedly appeared and disappeared. On each stream, one item rotated and one item changed from appearance to appearance. Infants preferred the streams in which the cued item (i.e., rotated) changed color to streams in which the cued item did not change color. The same results were replicated in eye-tracking change detection studies. For example, Cantrell et al. (2019) used a one-shot change detection task with two items, and the two items onset at different times (i.e., one object appeared briefly before the other object appeared). Six-month-old infants detected a change during the test array if they previously fixated the cued to-be-changed item during the sample array. Thus, infants store items they attend to in VSTM.

However, this ability appears to develop during infancy. In a one-shot change detection task, Beckner et al. (2020) induced infants to fixate an item during the sample by rotating that item. Eight-month-old infants preferred the changed item if they previously fixated the to-be-changed item in the sample array, whereas 4-month-old infants equally preferred the changed item and non-changed item. Recall that Ross-Sheehy and Eschman (2019) found that only when 5-, 8, and 11-month-old infants, and adults fixated the to-be-changed item during the sample array, their change preference was above chance. They also found that 5-month-old infants detected changes at a smaller set size than did older infants. In a follow-up study, Eschman and Ross-Sheehy (2023) found that 5- and 8-month-old infants detected a change for at least two fixated items during the sample, although their change preference was higher when the last item they fixated (during the sample period) changed than when the second-to-last fixated item (during the sample period) changed.

In our study, we used infants' spontaneous looks to the to-be-changed item during the prechange period and replicated the previous findings (Eschman & Ross-Sheehy, 2023; Ross-Sheehy & Eshman, 2019). Thus, our findings complement previous work in which infants' looking during the prechange period was manipulated by making an item more salient (Beckner et al., 2020; Cantrell et al, 2019). We confirmed that only when 5- to 12-month-old infants looked at the to-be-changed item during the prechange period did they show a robust preference for the changed item during the postchange period. When infants did not fixate the to-be-changed item during the prechange period, infants' memory for the changed item were at chance.

It should be noted that our prechange period was short (800 ms including delay), and as a result most infants only sampled one item during the prechange period. This contrasts to the 1000 ms plus 500 ms delay of Ross-Sheehy and Eschman (2019) and the 1500 ms plus 500 ms delay of Eschman and Ross-Sheehy (2023). In both of these previous studies, infants were allowed more time to look at multiple items in the sample array, and as a result Eschman and Ross-Sheehy (2023) could examine infants' change preference for different fixated items (e.g., the last versus the second to last). As a result of our relatively short sample array, in our study, infants could only fixate one item. Nevertheless, our results suggest that when these 5- to 12-month-old infants viewed our arrays, they could effectively reduce them to a single item, or set size 1, and they preferred the change item when that was the item they encoded in VSTM.

4.3 Conclusion

One conclusion that might be drawn is that infants have a VSTM capacity of 3 or 4 items. In this and other studies (Eschman & Ross-Sheehy, 2023; Ross-Sheehy & Eschman, 2019), infants detected changes in arrays of 3 or 4 items. However, we (and others) have found that this change detection is specific to trials in which infants fixated the to-be-changed item during the prechange period (Beckner et al., 2020; Cantrell et al., 2019; Eschman & Ross-Sheehy, 2023; Ross-Sheehy & Eschman, 2019; Ross-Sheehy et al., 2011). That is to say that when infants did not fixate the to-be-changed item during the prechange period, they showed no memory for the changed item. Moreover, the infants in our study fixated only one item during the prechange period. Thus, our results show that when infants are presented with arrays of 3 or 4 items, they can detect a change in the one item they fixate during the prechange period.

We did not find evidence of an age effect in 5- to 12-month-old infants' change localization at set sizes 3 and 4. This contrasts with other work showing, for example, that younger infants detect changes at smaller set sizes than older infants (Beckner et al., 2020; Kwon et al., 2014; Ross-Sheehy et al., 2003; Ross-Sheehy & Eschman, 2019). Moreover, it is still unclear why 5- to 12-month-old infants in our study–as well as other studies (Eschman & Ross-Sheehy, 2023; Ross-Sheehy & Eschman, 2019) detect a change in a single item at set sizes 3 and 4 when previous studies suggested that change detection for multi-item arrays do not emerge until 7 or 8-months of age (Kwon et al., 2014; Oakes et al., 2006, 2009, 2011, 2013; Ross-Sheehy et al., 2003). The literature on infants' VSTM capacity is inconsistent and estimating set sizes is complicated. Variations in testing and stimulus creation may be contributing reasons for the consistency. Nonetheless, these studies have added to our understanding that infants have VSTM, how VSTM works in infancy, and how infants store information in their VSTM.

Taken together, this study offers new insights into infants' VSTM by testing a wide age range and utilizing eye tracking to observe what infants pay attention to on the screen when they are shown arrays of 3 or 4 items. By testing multiple set sizes, we were able to see the interaction between age and fixated to the to-be-changed item during prechange period on change preference at set size 4. Although we are unable to explain why there is no significant interaction between age and what infants fixated during the prechange period for set size 3, we do know that younger infants' change preference is more sensitive to what they paid attention to during the prechange period for set size 4. As for older infants, what infants paid attention to during the prechange period had little effect on their change preference at set size 4.

References

Beckner, A. G., Cantrell, L. M., DeBolt, M. C., Martinez, M., Luck, S. J., & Oakes, L. M. (2020). Visual short-term memory for overtly attended objects during infancy. *Infancy: The Official Journal of the International Society on Infant Studies*, *25*(3), 347–370. https://doi.org/10.1111/infa.12332

Buss, A. T., Ross-Sheehy, S., & Reynolds, G. D. (2018). Visual working memory in early development: a developmental cognitive neuroscience perspective. *Journal of Neurophysiology*, *120*(4), 1472–1483. https://doi.org/10.1152/jn.00087.2018

Cantrell, L. M., Kanjlia, S., Harrison, M., Luck, S. J., & Oakes, L. M. (2019). Cues to individuation facilitate 6-month-old infants' visual short-term memory. *Developmental Psychology*, *55*(5), 905–919. https://doi.org/10.1037/dev0000683

Chapman, R. S., & Hesketh, L. J. (2001). Language, cognition, and short-term memory in individuals with Down syndrome. *Down's Syndrome, Research and Practice: The Journal of the Sarah Duffen Centre / University of Portsmouth*, *7*(1), 1–7. https://doi.org/10.3104/reviews.108

Coffman, B. A., Murphy, T. K., Haas, G., Olson, C., Cho, R., Ghuman, A. S., & Salisbury, D. F. (2020). Lateralized evoked responses in parietal cortex demonstrate visual short-term memory deficits in first-episode schizophrenia. *Journal of Psychiatric Research*, *130*, 292–299. https://doi.org/10.1016/j.jpsychires.2020.07.036

Dovis, S., Van der Oord, S., Wiers, R. W., & Prins, P. J. M. (2013). What part of working memory is not working in ADHD? Short-term memory, the central executive and effects of reinforcement. *Journal of Abnormal Child Psychology*, *41*(6), 901–917. https://doi.org/10.1007/s10802-013-9729-9

Eschman, B., & Ross-Sheehy, S. (2023). Visual Short-Term Memory Persists Across Multiple Fixations: An n-Back Approach to Quantifying Capacity in Infants and Adults. *Psychological Science*, *34*(3), 370–383. https://doi.org/10.1177/09567976221136509

Hyun, J. S., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(4), 1140–1160. https://doi.org/10.1037/a0015019

Johnson, J. S., Spencer, J.P., Luck, S.J., & Schöner, G. (2009a). A dynamic neural field model of visual working memory and change detection. *Psychological Science*, 20, 568–577. https://doi.org/10.1111/j.1467-9280.2009.02329.x

Johnson, J. S., Spencer, J.P., & Schöner, G. (2009c). Moving to higher ground: the dynamic field theory and the dynamics of visual cognition. *New Ideas in Psychology*, 26, 227–251. https://doi.org/10.1016/j.newideapsych.2007.07.007

Káldy, Z., & Leslie, A. M. (2005). A memory span of one? Object identification in 6.5-month-old infants. *Cognition*, *97*(2), 153–177. https://doi.org/10.1016/j.cognition.2004.09.009

Kibbe, M. M., & Feigenson, L. (2016). Infants use temporal regularities to chunk objects in memory. *Cognition*, *146*, 251–263. https://doi.org/10.1016/j.cognition.2015.09.022

Kwon, M. K., Luck, S. J., & Oakes, L. M. (2014). Visual short-term memory for complex objects in 6- and 8-month-old infants. *Child Development*, *85*(2), 564–577. https://doi.org/10.1111/cdev.12161

Lange-Küttner, C. (2008). When One Is Really Two: Switching from Object to Place Memory in the A-Not-B search task. *International Journal of Developmental Science*, *2*(4), 370–386. https://doi.org/10.3233/dev-2008-2403

Leclercq, A.-L., Maillart, C., Pauquay, S., & Majerus, S. (2012). The impact of visual complexity on visual short-term memory in children with specific language impairment. *Journal of the International Neuropsychological Society: JINS*, *18*(3), 501–510. https://doi.org/10.1017/S1355617712000021

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

Luck, S. J. (2007) Visual short term memory. *Scholarpedia*, 2(6):3328.

Mitsven, S. G., Cantrell, L. M., Luck, S. J., & Oakes, L. M. (2018). Visual short-term memory guides infants' visual attention. *Cognition*, *177*, 189–197. https://doi.org/10.1016/j.cognition.2018.04.016

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: is number one of them? *Psychological Bulletin*, *128*(2), 278–294. https://doi.org/10.1037/0033-2909.128.2.278

Moher, M., Tuerk, A. S., & Feigenson, L. (2012). Seven-month-old infants chunk items in memory. *Journal of Experimental Child Psychology*, *112*(4), 361–377. https://doi.org/10.1016/j.jecp.2012.03.007

Oakes, L. M., Baumgartner, H. A., Kanjlia, S., & Luck, S. J. (2017). An eye tracking investigation of color-location binding in infants' visual short-term memory. *Infancy: The Official Journal of the International Society on Infant Studies*, *22*(5), 584–607. https://doi.org/10.1111/infa.12184

Oakes, L. M., Baumgartner, H. A., Barrett, F. S., Messenger, I. M., & Luck, S. J. (2013). Developmental changes in visual short-term memory in infancy: evidence from eye-tracking. *Frontiers in Psychology*, *4*, 697. https://doi.org/10.3389/fpsyg.2013.00697

Oakes, L. M., Hurley, K. B., Ross-Sheehy, S., & Luck, S. J. (2011). Developmental changes in infants' visual short-term memory for location. *Cognition*, *118*(3), 293–305. https://doi.org/10.1016/j.cognition.2010.11.007

Oakes, L. M., Messenger, I. M., Ross-Sheehy, S., & Luck, S. J. (2009). New evidence for rapid development of color-location binding in infants' visual short-term memory. *Visual Cognition*, *17*(1-2), 67–82. https://doi.org/10.1080/13506280802151480

Oakes, L. M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science*, *17*(9), 781–787. https://doi.org/10.1111/j.1467-9280.2006.01782.x

Parra, M. A., Abrahams, S., Logie, R. H., Méndez, L. G., Lopera, F., & Della Sala, S. (2010). Visual short-term memory binding deficits in familial Alzheimer's disease. *Brain: A Journal of Neurology*, *133*(9), 2702–2713. https://doi.org/10.1093/brain/awq148

Perone, S., Simmering, V. R., & Spencer, J. P. (2011). Stronger neural dynamics capture changes in infants' visual working memory capacity over development. *Developmental Science*, *14*(6), 1379–1392. https://doi.org/10.1111/j.1467-7687.2011.01083.x

Ross-Sheehy, S., & Eschman, B. (2019). Assessing visual STM in infants and adults: eye movements and pupil dynamics reflect memory maintenance. *Visual Cognition*, *27*(1), 78–92. https://doi.org/10.1080/13506285.2019.1600089

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2011). Exogenous attention influences visual short-term memory in infants. *Developmental Science*, *14*(3), 490–501. https://doi.org/10.1111/j.1467-7687.2010.00992.x

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*(6), 1807–1822. https://doi.org/10.1046/j.1467-8624.2003.00639.x

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, *111*(4), 695–707. https://doi.org/10.1016/j.jecp.2011.10.007